DOCUMENT RESUME

ED 198 529 CS 206 109

AUTHOR Anderson, Pamela L.
TITLE Cohesion as an Index for Written and Oral Composition of ESL Learners.
PUB DATE Nov 80
NOTE 15p.: Paper presented at the Annual Meeting of the Midwest Modern Language Association (22nd, Minneapolis, MN, November 6-8, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adult Students: Applied Linguistics: *Discourse Analysis: *English (Second Language): Evaluation Methods: Higher Education: *Language Proficiency: *Language Research: Language Skills: *Second Language Learning: Student Evaluation: *Verbal Communication: Writing Research
IDENTIFIERS *Cohesion (Written Composition): Holistic Scoring

ABSTRACT

To determine the usefulness of cohesion frequency analysis in differentiating proficiency in the oral and written composition of English as a second language (ESL) learners, 22 adult learners of ESL were asked to write and tell a story. After transcription, the data were analyzed by holistic evaluation, cohesion form and frequency count, and error analysis of cohesive elements used. Analysis showed all five forms of cohesion (reference, substitution, ellipsis, conjunction, and lexical) were used, although each subject did not necessarily use all forms. From a statistical perspective, cohesion frequency analysis of the five separate forms related inconsistently with the holistic evaluation results and only once with the teaching of English as foreign language scores. Total cohesion frequency related to neither regardless of mode. Cohesion analysis on a frequency basis produced little usable information. However, if frequency and quality of the cohesive devices used could be measured, it is probable that applied linguists could learn more about how ESL learners acquire and use intersentence connectors. (HOD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED198529

Cohesion as an Index for Written and Oral Composition of ESL Learners

Paper presented at
MMLA, Linguistics I
November 6 - 8, 1980

Pamela L. Anderson
University of Hawaii

Determination of proficiency in composition has been accomplished by use of various measurement indices which generally reflect the focus of language theory. Since the 1960's linguists have considered the sentence of particular importance for description and explanation, a consideration which has been reflected by applied linguists in analysis procedures. For example, compositions have been analyzed by T-units (Hunt, 1965) or by modified versions of this unit such as the error-free T-unit (Scott & Tucker, 1974). These units have been considered adequate at the sentence level by most investigators. Recently the focus of many linguists has shifted beyond the sentence level. Sentence level indices, therefore, are no longer adequate. One potential index which relates sentences to one another is derived from Halliday and Hasan's conception of cohesion in English (1976). This index, as codified by Halliday and Hasan, has been applied to both native and non-native learners' written language. One objective of the investigation reported here was to determine the usefulness of cohesion frequency analysis in differentiating

CS206109

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Pamela L. Anderson
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U S DEPARTMENT OF HEALTH EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

proficiency in the oral and written composition of English as a Second/Foreign Language (hereafter ESL) learners. The cohesion analysis was related to the learners' TOEFL scores and holistic evaluations of their compositions to consider this potential.

Since this investigation was designed to consider a new index for ESL data, the language type elicited was limited to narration. Thus all results from the cohesion analysis are similarly limited, as other types of language might result in different amounts and forms of cohesion.

Cohesion, as defined by Halliday and Hasan, is that which unifies a text or makes it coherent with itself. This is achieved by using specific devices which relate ideas and structures across sentence boundaries. The text, or group of related sentences, must also be coherent with respect to the situation or context; however, this is termed register not cohesion by Halliday and Hasan. Cohesion may be subdivided into types and forms.

The types of cohesion are determined by location of the antecedent to which the cohesive device points. A cohesive device pointing to an antecedent within the context but not mentioned in the actual text is EXOPHORIC, as is "that" in example one, Table 1. Exophora is not considered in analyzing textual cohesion as it extends beyond the text. With ESL data, however,

it is important to recognize that exophoric cohesion may lie behind many "errors." The learner may not recognize that a non-textual antecedent needs to be clarified in the text before it is alluded to further.

ENDOPHORA is cohesion within the text. An endophoric cohesive device may either point forward (CATAPHORA) or backward (ANAPHORA). Anaphora is the most commonly found type of cohesion while cataphora is rarely cohesive, since it tends to point forward within the bounds of the sentence. Examples of both anaphora and cataphora as cohesive devices are given in Table 1.

Table 1

1.	EXOPHORA		That is my house on the corner.
2.	ENDOPHORA	A. ANAPHORA	Those children are in danger. The little one just fell.
		B. CATAPHORA	This is how to get back. Take 7th street west ....

Types of Cohesion

Cohesion, as mentioned above, may also be divided into forms. The five forms of cohesive devices categorized by Halliday and Hasan are reference, substitution, ellipsis, conjunction, and lexical cohesion. All five contain devices which point to another part of the text in order to create unity. As each is considered, please refer to Appendix A for examples.

Reference is a meaning relationship between s which names one of three relationships: a perso onominal), the proximity of a location (demonstrative), or t dentity or similarity of two items (comparison). Substi n, unlike reference, is an item rather than a meani relationship in that the substitute form fulfill structurally the same function as the antecedent. Substitution may occur fo a nominal, a verbal, clausal constituent. Ellipsis is similar to sub tion except the substituted form is empty. Ellipsis also occur for nominals, verbals, or clauses. T three forms of cohesion are primarily structural eas conjunction and lexical cohesion are primarily s ic.

Conjunction is unlike the previous three co n forms since the device, or the conjunction (conj ve phrase), used to indicate the relationship does simply refer to or replace another item in the d rse. A conjunction is merely a device which points ou strengthens, and clarifies the semantic relation between two sentences. In the Example "They wer unhappy. Nonetheless, they smiled pleasantly to us a safe journey." it is the cohesiveness betwe e internal state of unhapiness and the external st f smiling pleasantly which is illuminated by th ction, "nonetheless." Halliday and Hasan enumerate fo a- tionships shown by conjunctions. These are addi adversative, causal, and temporal relationships.

The final form of cohesion is lexical cohesion. This refers to the properties of words which create unity by pointing out relationships or reminding the reader/listener of other portions within the discourse. Halliday and Hasan state that lexical cohesion results from a selection process which includes repeating vocabulary, using synonyms, using superordinates, using general terms for specific ones, and using collocation.

Cohesion may also be attained through other means such as sentence structure and progression of verb times/tenses throughout a passage. Halliday and Hasan recognize this fact, but do not attempt to codify these in their analysis. Others have considered various protions of cohesion in their work, such as Loban (1963) who discusses use of substitute forms, yet the scheme described by Halliday and Hasan has the advantage of being relatively concise, including most of the cohesive forms used in language, and being quantifiable.

Since the publication of Cohesion in English, the scheme suggested by Halliday and Hasan has been used to analyze the written language of skilled authors, to analyze reading comprehension of first language learners, and to analyze listening comprehension. Analysis of freshman composition by cohesion is currently being undertaken. Thus far, however, no patterns for use of cohesion by native speakers have been identified.

In second language research, a few investigations of cohesion usage have been reported in the last two years. Again, patterns have not been established. Thus, this investigation is a continuation of preliminary research into the use of cohesion as an index for analyzing the language of ESL learners.

The study explored the written and oral narrative language of 22 adult learners of English as a Second or Foreign Language. To elicit comparable samples of oral and written composition, half of the subjects were asked to tell then write a story while the other half performed the tasks in reverse order. The abilities of the two groups were equivalent according to results of a T-test using TOEFL scores as the index. The order of tasks performed did not affect results in language. After transcription, the data were analyzed by holistic evaluation, cohesion form and frequency count, and error analysis of cohesive elements used. The cohesion analysis was statistically related to holistic evaluation results and TOEFL scores. Non-linguistic parameters were controlled statistically for relationships with the language produced. Non-parametric statistics for small sample sizes were used, and the level of confidence was set at the 0.05 level throughout.

The analysis of the subjects' language showed that all five forms of cohesion appeared in both oral and written composition, although each subject did not

necessarily use all forms. The average number of cohesion devices attempted in oral composition was 55.27, which was significantly higher than that of written composition at 42.96. Of those devices attempted, an average of 46.18 were used correctly in oral language and 37.82 were used correctly in written language.

The forms of cohesion averaged in frequency in the following manner:

Oral reference	40.73
Written reference	31.82
Oral substitution	1.10
Written substitution	1.40
Oral ellipsis	2.50
Written ellipsis	2.00
Oral conjunction	5.86
Written conjunction	2.33
Oral lexical	5.08
Written lexical	5.91

These averages include only examples which could be clearly identified as an attempt at the intended form, whether or not that attempt was successful. One problem with cohesion analysis for ESL data is exact determination of the intended form and whether or not its function is cohesive.

The relationship between the cohesion frequencies and the results of the holistic evaluations were not significant overall. Within the separate forms of cohesion, however, three were significant. Frequency of reference cohesion in written language and the holistic evaluation results of the written compositions were significantly related. The relationship showed a negative

correlation coefficient between the holistic evaluations and frequency of correct reference cohesion. In frequency of conjunction usage the holistic evaluation of the oral language was related to both frequency of conjunction in oral and written language. The correlation coefficient was negative for frequency of correct conjunction usage in oral language, but it was positive for frequency of correct conjunction usage in written language. These findings were unsettling, since a Wilcoxon Matched-Pairs Ranked-Signs test had shown that raters of the compositions (holistically) viewed oral and written language similarly. The three raters were also shown to be reliable among themselves on a Friedman Two-way Analysis of Variance. Yet, in viewing conjunction cohesion, it seems that raters did not view the two modes similarly. The inconsistencies in these results cannot be explained adequately here; possibly they result from the small sample size, but further research is recommended.

In addition to relating cohesion to holistic evaluation results, I considered the TOEFL as an index of the subjects' proficiency and related their scores to the cohesion analysis via Spearman Correlation Coefficient tests. Like the holistic evaluation results, the TOEFL scores showed no significant relationship to total cohesion frequency for either oral or written language. When the forms were separated, the TOEFL

did show a significant relationship to the frequency of reference used in writing. This coefficient was positively related to correct usage of reference, a fact which is not surprising since the TOEFL tests use of pronouns and agreement, use of demonstratives and the definite article, and use of comparative adjectives.

From the above purely statistical perspective, cohesion frequency analysis of the five separate forms relates inconsistently with the holistic evaluation results and only once with the TOEFL. Total cohesion frequency relates to neither regardless of mode. Thus, according to the dimensions of this particular study, cohesion frequency seems to muddle the picture as it relates to a commonly accepted measure of unity in composition and to a similarly accepted measure of English proficiency for non-native learners. On the other hand, cohesion frequency does show significant differences between oral and written narrative for these subjects.

Despite the inconclusive results of the statistical analysis, several trends were seen and are providing bases for further research. For example, there may be a hierarchy using cohesion forms by non-native learners. Most ESL learners seem to rely on lexical cohesion, although this does not seem to be a selection process, but a result of limited vocabulary. One learner wrote (with spelling corrections):

> it was snowing and the snow it snowed very hard. The harder it came the worst the light become. The light is not good light enough to see and the snow snowed more and my car it was stuck there.

A more sophisticated writer would not have made the same selection of vocabulary. Yet, this subject did not seem to have alternate vocabulary to state the ideas and so used repetition to create a cohesive, if tedious, passage. Whether or not this can be counted toward a cohesion analysis is debatable since the choice of lexical items was probably not governed by a desire to unify the passage, but merely to express the content.

The second most used form of cohesion was reference. Very little difference in frequency of reference use was detected between written and oral composition. The most used form of reference was pronominal, followed by the forms of demonstrative then comparison. This was true for both modes of language. Differences also did not occur between modes in the amount of correct and incorrect usage. Were there differences in use of reference between oral and written composition? After continued investigation, results suggest tentatively that the differences exist but not in frequency. Some differences exist in location of antecedent, with oral language referring more often to non-textual antecedents

(i.e., using more exophoric cohesion) than written language does. Again, these results are tentative, but use of reference cohesion by second language learners may prove illuminating when form rather than frequency is considered.

Further observations show that substitution and ellipsis appear rather infrequently. Yet, when substitute forms are used or ellipsis was seen, they were nearly always correct. In addition, these two forms were present only twice in oral data when not also in the same subject's written data. The reverse occurred more frequently. Furthermore, subjects using the forms of substitution and ellipsis did not fall exclusively into the top one-third or the bottom one-third of the sample on either rater holistic evaluation results or the TOEFL scores. Thus, substitution and ellipsis do not appear to be "advanced" forms of cohesion, but when they appear they are used correctly. With further research, it may be postulated, a difference in acquisition of these forms for oral and written language could be discovered.

Like substitution and ellipsis, conjunction was also used relatively infrequently. Most of the conjunctions used in both oral and written composition were additive. In written composition the majority of these were used correctly, while in oral composition this was not the case. The difference in appropriateness in using

conjunctions also deserves further investigation. Since conjunct cohesion results related to the holistic evaluation results were inconclusive in a frequency analysis, investigation into variation of items or quality when using conjunctions seems warranted.

To what conclusions about cohesion analysis does this investigation lead? In considering the language used in composition by ESL learners, whether oral or written, applied linguists have recognized the need for indices beyond the sentence level. One such index is cohesion analysis as proposed by Halliday and Hasan, although they did not use second language data per se. Although parts of the Halliday and Hasan scheme have been used by other researchers, their scheme and description is the most complete, concise, and quantifiable format for measuring intersentence connectors available today. This investigation shows that mere quantifying of cohesion forms may differentiate between oral and written language of ESL learners, but does not differentiate proficiency levels as indicated by those learners' TOEFL scores. Nor does total cohesion quantity seem to be a factor considered systematically in holistic evaluation of compositions.

Cohesion analysis on a frequency basis leaves us with little usable information. This does not mean that cohesion analysis is useless for oral and written composition of ESL learners. If frequency and quality

of the cohesive devices used could be measured, it is probable that applied linguists could learn more about how ESL learners acquire and use intersentence connectors. However, this one index alone, just as any other index used alone, cannot give an overall picture of the learners' language.

References:

Halliday, M.A.K. and R. Hasan. Cohesion in English. London: Longman, 1976.

Hunt, K.W. Grammatical structures written at three grade levels. Champaign, Illinois: NCTE Research Report No. 3, 1965.

Loban, W.D. The language of elementary school children: A study of the use and control of language effectiveness in communication, and the relations among speaking, reading, writing, and listening. Champaign, Illinois: NCTE Research Report No. 1, 1963.

Scott, M.S. and G.R. Tucker. Error analysis and English language strategies of Arab students. Language Learning, 1974, 24, 69-97.

Appendix

Forms of Cohesion

Reference	
Pronominal	John wanted a Camero. Then he saw the TV commercial about Datson.
Demonstrative	The oak table is too large; the maple one is too small. Well, I will take the maple one.
Comparative	The other table just would not fit in the room.
Substitution	
Nominal	This one is just right.
Verbal	John went to the store. So did I.
Clausal	Did you run the car into that pole? I did not!
Ellipsis	
Nominal	I bought the huge yellow basket. The yellow? I thought the red?
Verbal	I could not have said more. Yes, I could.
Clausal	Have you taken the cat outside? Yes.
Conjunction	
Additive	I hate school. And I hate homework!
Adversative	John has been up for 72 hours. Nonetheless, he looks great!
Causal	I was feeling terrible. So, I left.
Temporal	She flew up the wall. Then she escaped out the window.
Lexical	
Repetition	I saw the cat. That was the biggest cat I've ever seen!
Synonomy	The man was huge. Then this giant ....
Superordination	Marie wanted a pet. Her grandmother brought a puppy over for her birthday.
General item	The thing is, Marie is allergic to dogs.
Collocation	They were in heaven. But, as always, hell was not far away.